THE COMPARISON OF SELF-ADMINISTERED, UNSUPERVISED AT HOME TESTING AND AT SCHOOL, SUPERVISED GROUP TESTING OF THE IMPACT BASELINE COMPOSITE SCORES IN ONE INTERSCHOLASTIC HIGH SCHOOL

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

KINESIOLOGY AND REHABILITATIVE SCIENCE

August 2021

BY

Darian Brothers, LAT, ATC

Thesis Committee:

Nathan Murata, Chairperson
Christopher Stickley
Kyoko Shirahata

Keywords: Mild traumatic brain injury, concussion assessment, high school student athletes, computer-based neuropsychological test, Immediate Post-Concussion Assessment and Cognitive Testing, baseline, group administration, self-administered

Abstract

Context: Baseline concussion testing has been a common practice in head injury management, and the administration environment suggests to have an effect on the outcome scores. Due to the COVID-19 pandemic, social distancing restrictions prohibited concussion baseline testing in group settings. **Objective:** To compare Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) baseline scores of supervised groups to baselines taken unsupervised individually. **Design:** Retrospective study. **Participants:** 50 different individuals were selected from both the 2019-2020 and 2020-2021 school years; 2019-2020 took the baseline in a supervised group setting, 2020-2021 were unsupervised and self-administered the baseline. Both groups were from the same institution, matched by age, gender, sport, and history of concussions. **Procedures:** The 2019-2020 athletes took the ImPACT baseline in the institute's library, supervised by the employed athletic trainers. The 2020-2021 athletes were given an institution specific access code to input in order to self-administer the baseline at home (or where they had computer access). MANOVA was used to compare the composite scores (verbal memory, visual memory, visual motor speed, reaction time, and impulse control) and an Independent T-test was used to compare the symptom scale. **Results:** No significant difference was found between the composite scores of the two groups [V = 0.45, F(5,94) = 0.89, *p* = .49]. One symptom score, visual problems, was deemed significantly higher in the supervised group. **Conclusion:** The influence of certain test administration factors can have an effect on testing outcomes for every individual and must be considered when administering concussion baseline testing. Consistent control of administration and scripted instructions in future studies will fill the gaps that remain in the debate of baseline administration.

Word Count: 270

Table of Contents

List of Figures and Tables

List of Abbreviations

Immediate Post Concussion and Cognitive Testing---ImPACT

Neuropsychological---NP

Return to Play--RTP

Attention Deficit Hyperactivity Disorder---ADHD

Hawaii High School Athletic Association--HHSAA

Learning Disability--LD

Post-Concussion Symptom Scale---PCSS

Mild Traumatic Brain Injury---mTBI

Reaction Time--RT

Reliable Change Index---RCI

Multivariate Analysis of Variance---MANOVA

Centers for Disease Control and Prevention--CDC

Intraclass Correlation Coefficient--ICC

Athletic Trainer---AT

Introduction

Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) is a commonly used computer-based neuropsychological (NP) test battery that is used as a tool for healthcare professionals when assessing possible concussions.[1] ImPACT testing requires the user to input demographic information (e.g., age, sex), and to report any consistent concussion related symptoms. ImPACT testing covers six modules: Word Memory, Design Memory, X's and O's, Symbol Matching, Color Match, and Three Letter Memory. These six modules record several scores that reflect the testers speed and accuracy.

Based on the module scores, five composite scores of verbal memory (memorizing words), visual memory (memorizing images), visual motor speed (time it takes to process visual stimuli), reaction time (time it takes between a stimulus and response), and impulse control (the ability to quickly answer correctly) are generated to assess one's cognitive function and establish a baseline. For example, visual motor speed outputs its score by taking the speed of the correct responses of X's and O's and Three Letter Memory modules. A couple studies point out how visual motor speed and RT scores can be influenced by the type of test administration the tester is subjected to, such as being taken in a group or taken individually.[2 3]

Baseline measures are helpful because the scores can be compared with the same individual's post-concussion scores to assist healthcare professionals in the return to play (RTP) decision making process. A baseline measurement is useful in ascertaining a student athlete's current neurocognitive state; however, this baseline measurement may not be a true reflection of someone's cognitive best. Cognitive best in this case means a reflection of the user's mental speed and accuracy based on the testing module and composite scores.

There are variables that influence a person’s overall ImPACT score including but not limited to; motivation, fatigue, and pre-existing factors such as age, sex, history of Attention Deficit Hyperactivity Disorder (ADHD), and previous concussions.[3-6] Another potential common threat pertinent to a baseline measurement is called “sandbagging.”[1] Sandbagging means the user intentionally tried to achieve a low score, which connotes an invalid test whereby the user scored below normal. Normal in this context pertains to the average scores of over 16,000 tests taken from those aging 12-59 years old. ImPACT separated the collected tests by age and sex and the normative data scores (mean of the scores) are dependent on the users specific age and gender. ImPACT has set five “red flags” to help identify individuals who are suspected of sandbagging, as well as invalidity indicators that reveal an invalid test.[1 7] Invalid tests may have been influenced by some pre-existing factors associated with the user and not necessarily be contributed to a concussion.[1]

Testing environment (testing in groups or individually) and the presence of test proctors can possibly influence composite scores and rate of invalid tests. However, these topics need further discussion since only a few studies have been conducted using group proctors. Factors such as testing environment, convenience, supervision, and time restraints influence ImPACT baseline testing and therefore most assessors complete baseline testing in a proctored, group setting to complete multiple tests in a timely manner. Moser et al. concluded that group settings had higher invalid scores compared to individual test-taking[2] while French et al. and Vaughan et al. found no differences between the groups and concluded that the results were due to the subjects having proper instructions prior to testing.[8 9] Kuhn et al. compared proctored testing to unsupervised self-administered (not proctored) testing.[3] The study found that the group setting had higher visual motor speeds and faster reaction times but no differences in test validity;

concluding that the unsupervised individuals assumed the test only scored accuracy as opposed to speed and reaction times as well.

Due to the world-wide pandemic of the COVID-19 virus, the Hawai‘i State Department of Education (HIDOE) provided education by “distance learning” format during the 2020-2021 school year. In addition, the Hawai‘i High School Athletic Association (HHSAA) suspended all school affiliated sports at the start of the Fall 2020 school year.[10 11] The standard concussion management protocol for secondary school recommends school athletic trainers to conduct ImPACT baseline testing prior to the season of 9th and 11th grade. Given the protocol, most athletic trainers typically administered ImPACT baseline in group settings; however, it was difficult to continue group ImPACT baseline testing due to the unexpected challenges and restrictions. As a result, some high school institutions in Hawai‘i chose to let their athletes take the ImPACT baseline at home, in hopes of a speedy transition when sports started to take effect. The occurrence of another wave of COVID-19 has resurfaced and the uncertainty of group ImPACT testing remains in the near future. To this end, the purpose of this retrospective study was to compare proctored (by an athletic trainer [AT]), group ImPACT baseline composite scores (prior to the pandemic) to at home, unsupervised by an AT, self-administered (during the pandemic) scores. This comparison could help healthcare professionals make an informed decision when it comes to the administration of ImPACT baselines if the restrictions set by COVID-19 continue.

Hypothesis

Based on previous research that looked at the comparison of group testing, individual testing, proctored and un-proctored testing separately, it was hypothesized that there will be a difference between the groups’ composite scores.[2 3 8 9] More specifically, self-administered, at-home tests will have slower reaction time and visual processing speed composite scores than the baselines taken in a supervised group.

Methods

This retrospective study investigated the differences in ImPACT baseline composite scores from one interscholastic high school in two consecutive school years; one administered at school in a supervised group setting versus unsupervised, self-administered at home. All participants had consent forms signed by their guardians. Internal review board approval from the University of Hawaiʻi Committee of Human Studies (#18431) allowed access and use of all ImPACT scores from Interscholastic high school student athletes in the state of Hawaiʻi.

High School's ImPACT Baseline Procedure

The participating private high school institution administered ImPACT to their students every two years, in the beginning of 9th and 11th grade season. If an athlete sustained a concussion, the institution would have the athlete take a post-injury ImPACT test within 72 hours of injury and again when they were symptom free. The school's athletic trainer reviewed every post-injury test and makes clinical decisions based on the score differences from baseline. These decisions included sending the scores to a neuropsychologist, who was trained in ImPACT interpretation, when it deems fit.

Participants who were administered ImPACT in the 2019-2020 school year would sign up for a designated time and date with the schools AT to take the baseline. The baseline schedules were dependent on the availability of the school's library and the student athletes. The student athletes would come to the library at the time and date they signed up for. The library was equipped with enough desktop computers to accommodate about 20 student athletes to take the baseline simultaneously. Once the student athletes entered the access code into the ImPACT website, they could begin the baseline and when they were finished they could leave. Manual stated that test administration takes about 20-25 minutes; however, individuals age and speed in responding are dependent.

Due to restrictions of having large groups accumulate together due to COVID-19 in the 2020-2021 school year, the private high school had their student athletes take the ImPACT baseline at home. The student athletes were provided with the ImPACT website and an access code that was linked to their institution and took the baseline, unsupervised by an AT, from their residence (or any place they had internet access). The students took the baseline in their free time, so there was no set date or time they were taken. As a result of no supervision, it is unknown what device (i.e. iPad, desktop, laptop) was used and if an external mouse was utilized (recommended by ImPACT manual).[1] All of the student athletes scores (from both groups) were stored in the ImPACT database.

Participants

Participants were student athletes from a local private high school, ages 13-18y/o, from the 2019-2020 and 2020-2021 school years. The initial population for the 2019-2020 and 2020-2021 school years include 332 and 182 participants, respectively. Inclusion criteria were: (1) high school student athletes (educational level 9-12) and (2) English reported as their primary language. Exclusion criteria included: (1) reported a history of a learning disability (LD) or Attention Deficit Hyperactivity Disorder (ADHD). Due to the school's procedure of only doing baselines every two years, this means the two groups are different individuals but the cohorts are matched by sex, age sport, and number of previous concussions and participants were de-identified. De-identified data extraction included age, sex, sport, history of concussions, and the composite scores of verbal memory, visual memory, visual motor speed, reaction time, and impulse control.

Materials

The NP test battery used for the cognitive baselines was the ImPACT test (ImPACT Applications Inc.©, Coralville, IA), version 3. ImPACT was taken through their online website, so internet must be available at the time the test is administered. ImPACT had six testing modules that calculate five composite scores (verbal memory, visual memory, visual motor [processing] speed, reaction time, and impulse control). The five composite scores and the users input demographics was the material extracted and utilized. Based on previous studies, ImPACT's had a 6% rate of invalid baseline scores for those 10-18 years of age and an >80% chance of detecting potential sandbagging.[12] ImPACT's test-retest reliability for all composite scores fell between low and moderate, meaning it does not achieve good reliability.[13] A systematic review concluded that ImPACT's diagnostic accuracy and predictive validity are inconclusive; however, it is shown that it yields a 94.6% sensitivity and 97.3% specificity in it's results.[1 14]

ImPACT Testing Procedures

In this section, it will go over specifically what both groups were asked and tested on when they started the ImPACT baseline test. Both groups took the same tests and were asked the same questions in the same order.

When accessing, ImPACT had the user input their specific login information. This login information was administered through a school affiliation, in which a specific code is given so their information is stored with that school affiliation data. Next, ImPACT asked the user to input their demographics. This included their sex, age, ethnicity, native language, educational level, input of any special needs (ADHD, dyslexia, autism, etc.), concussion history, sport, and if there is any relevant medical history to add. Next is the symptom scale, and finally the user selected either “Baseline” or “Post-injury” test option to start.

ImPACT showed 22 symptoms on their post-concussion symptom scale (PCSS) that are linked to a possible mild traumatic brain injury (mTBI), also known as a concussion. Here are some symptoms that are listed (but not limited to): headache, nausea, balance problems, dizziness, sensitivity to noise or light, sadness, difficulty falling asleep, and difficulty concentrating (see Appendix A for an entire list of ImPACTs PCSS). The user selected any symptoms they have experienced in the past 24 hours and rate them on a scale from zero to six. Zero being they have not had that symptom and six meaning the symptom is very severe. At the end, ImPACT generated a total symptom score and a symptom severity score. The total symptom score is the sum of all the symptoms the user is experiencing (if they put a one or above on the severity). For example, if the user put four for headache, three for nausea, and zero for the other 20 symptoms then their total symptom score would be two. The symptom severity score was a sum of the severity rate the user selected. So, if the user selected a severity of two on all 22 symptoms, their severity score would be 44. After the PCSS the user began the test battery.[1]

The NP portion had six modules: Word Memory, Design Memory, X’s and O’s, Symbol Matching, Color Match, and Three Letter Memory. Word Memory was used to evaluate verbal recognition memory and attention processes. This module showed the test taker 12 words, one at a time, for 750 milliseconds per word. It showed the list of 12 once more. Afterwards, the test ran through random words and had the user select “yes” or “no,” identifying if that particular word was one of the original 12. After the user completed all six modules, they were tested again to recall the 12 words from the beginning. Design Memory evaluated visual recognition memory and attentional processes. This module was the exact same as Word Memory, but instead of the user having to recall words the user had to recall a fractal design (see Appendix B for a visual of the Word Design Module). The user had to recall the designs after all the completion of the modules as well.

X's and O's module consisted of two parts: a distractor task and a memory task. This module was used to assess visual processing/visual motor speed, and visual working memory. The distractor task was given first and the user has to do a specific action of hitting the "Q" button on the keyboard when a red circle appears or hit the "P" button when a blue square appears. The user does not know which will appear, and therefore must recognize and hit the corresponding button as quickly as they can. After the distractor task, the user did the memory task. The user was shown a screen for 1.5 seconds in which a random assortment of X's and O's was presented. Three of the X's or O's on the screen will be yellow in color. After the 1.5 seconds the user was sent back to the distractor task, after they complete the distractor task the same screen of X's and O's will be presented. The user was instructed to recall which X's or O's were in yellow from the previous screen (see Appendix C for a visual of the X's and O's module). The user completed four trials of that sequence.

Symbol Matching would follow, evaluating memory, learning, and processing speed. The user was given a grid. The top of the grid had nine common symbols (square, arrow, circle, etc.) and directly under each symbol is a number button, one through nine. Below this grid one of the symbols will appear and the user has to click on the number button corresponding to that symbol. There are 27 trials of this. After the 27th trial all the symbols will be removed from the grid. The symbols appear below the grid and the user must recall the correct symbol to its corresponding number. Module five is Color Match, this will measure impulse control and response inhibition. The user will first be tested by clicking a blue, red, or green button as instructed to ensure that color blindness will not affect this task. In this module, the name of a color will be presented on the screen and the word will also be colored. The user must click the box as fast as they can when the word and the color match. For example, the user would click the screen if the word “red” was displayed in red ink and refrain from clicking the screen if the word “blue” was written in green ink.

The last module, Three Letters, measures visual-motor response speed and working memory and the user is first given a distractor task. A 5x5 grid with the numbers 1-25 is presented in a randomized order. The user must click on the numbered buttons as quickly as possible in backwards order, starting with 25. After the user is finished, they will be given three consonants to memorize. After the three letters disappear the randomized grid reappears, in a different order, and the user has 18 seconds to get as far as they can clicking the numbers in backwards order. After 18 seconds, the user had to type in the three letters that were presented to them prior. The user will complete five trials of this sequence.

ImPACT Module Scoring Procedure

Each module received its own score, consisting of six composite scores that are calculated after the user completes the test battery. Each module develops its own score to reflect the user's accuracy and speed. Word Memory identifies the number of correct “hits” and “distractors” that were immediately identified. “Hits” being the amount of times the user identified one of the 12 words given to memorize and “distractors” being the words used that were not part of the 12 (distractor words). The total of the two are then added together, divided by 24, and multiplied by 100 to give the “learning percent correct.” Then it identifies the amount of “hits” and “distractors” that were correct but the reaction was delayed. These two are added together, divided by 24, and multiplied by 100 to give the “delayed memory percent correct.”

The “learning percent” and the “delayed percent” are then added and divided by two to give the “total percent correct.” Design Memory score was calculated the exact same way.[1]

X’s and O’s score takes the “total Correct (memory)” from the amount of correctly identified X’s and O’s (12 max total). The rest of the scoring is based on the distractor test. “Total correct (interference)” and “total incorrect (interference)” gives the score of the number of correct and incorrect responses respectively. “Average correct Reaction Time (RT) (interference)” and “total incorrect RT (interference)” gives the average RT for the correct and incorrect responses respectively. Symbol Match calculates the “total correct (visible)” and “total correct RT (visible).” Taking the total amount of correctly identified symbols (27 max total) and the average reaction time of the correct responses. “Total correct (hidden)” and “total correct RT (hidden)” calculates the correct responses when the symbols were hidden.[1]

Color Match calculated “total correct,” “total commissions,” “average correct RT,” and “average commissions RT.” These show the number of correct and incorrect matches/responses and average reaction time for correct and incorrect matches/responses respectively. Three Letters calculates “total sequence correct,” which is the total number of letter sequences the user got correct (out of five). “Total letters correct” is the sum of correctly remembered letters, regardless of the sequence (15 max total). Based on the last score, “percent of total letter correct” is calculated. For the distractor portion, “average time to first click” is the time it took for the user to make the first click of the mouse (identifying number 25 of the backwards sequence). “Average counted” and “average counted correctly” are calculated based on the average number of numbers clicked (regardless of errors) and the average number of numbers clicked that were in the correct sequence respectively (see Appendix D for a blank clinical report of the module scoring).[1]

ImPACT Composite Scoring Procedure

These six modules produced five composite scores and used dependent variables: verbal memory, visual memory, visual motor speed, reaction time, and impulse control. Composite indices are expressed in percentiles and used to provide the information for the user's cognitive performance. Verbal memory is the average score of modules one, four, and six combined. So, the sum of module one's "total percent correct," module four's "total correct hidden" (divided by 9 and multiplied by 100), and module six's "percent of total letters correct." The sum of these three is then divided by three to get an average percentile, also known as the verbal memory composite score. Visual memory takes the average of module two's "total percent correct" and module three's "total correct (memory)" (divided by 12 and multiplied by 100) (see Appendix E for a visual of how visual memory composite score is calculated).[1]

Visual motor speed is the average of module three's "total correct (interference)" (divided by four) and module six's "average counted correctly" (multiplied by three). Reaction time composite is calculated by taking the average of module three's "average correct RT (interference)," module four's "average correct RT (visible)" (divided by three), and module five's "average correct RT." Impulse control composite is calculated by taking the average of module three's "total incorrect (interference)" and module fives's "total commissions."

Test takers who have both a baseline and a post-injury test from a mTBI, a RCI score was given for each module and composite score. The RCI score is there to help the healthcare professional identify when there is a statistically significant change between the user's baseline and post-injury test scores. If there is a significant difference the score will appear in red ink on the user's clinical report.

Statistical Analysis

All data were analyzed using Statistical Package for Social Science (SPSS) Version 28.0 (IBM®, Chicago, IL). Multivariate analysis of variance (MANOVA), with an alpha level of $p < 0.05$, was conducted to assess differences in composite scores between the two groups. Independent variables are identified as a supervised group and unsupervised self-administered group. Dependent variables were the five composite scores, which were verbal memory, visual memory, visual motor speed, reaction time, and impulse control. Based on the G*Power analysis[15], the estimated sample size for repeated measures (between factors) MANOVA would be 44 with effect size of 0.25 and power level at 0.95. The 22-symptom checklist were also analyzed using an independent-sample t-test, with an alpha level of $p < 0.05$.

Results

Data screening

A total of 528 ImPACT data were analyzed in this study. Of which, 346 were from at school administration, under supervision in 2019 and 182 were from at home, unsupervised administration in 2020. Then, the data was screened and any irrelevant data such as post-concussion ImPACT data (n=14), data with missing sports (n=79), out of target age range (n=67), and duplicate baseline (n=10) were removed. Then, samples from 2019 were matched with 2020 with the following demographic information: age, sex, sport, and history of concussion. After sample matching, there were a total of 104 students (at school n=52 & at home n=52). Descriptive analysis of the data revealed a total of two outliers, and Mahalanobis distance determined their removal. The removal of these outliers in the composite score finalized a total of 50 matched participants in each group. See Figure 1 for a visual of the data screening.

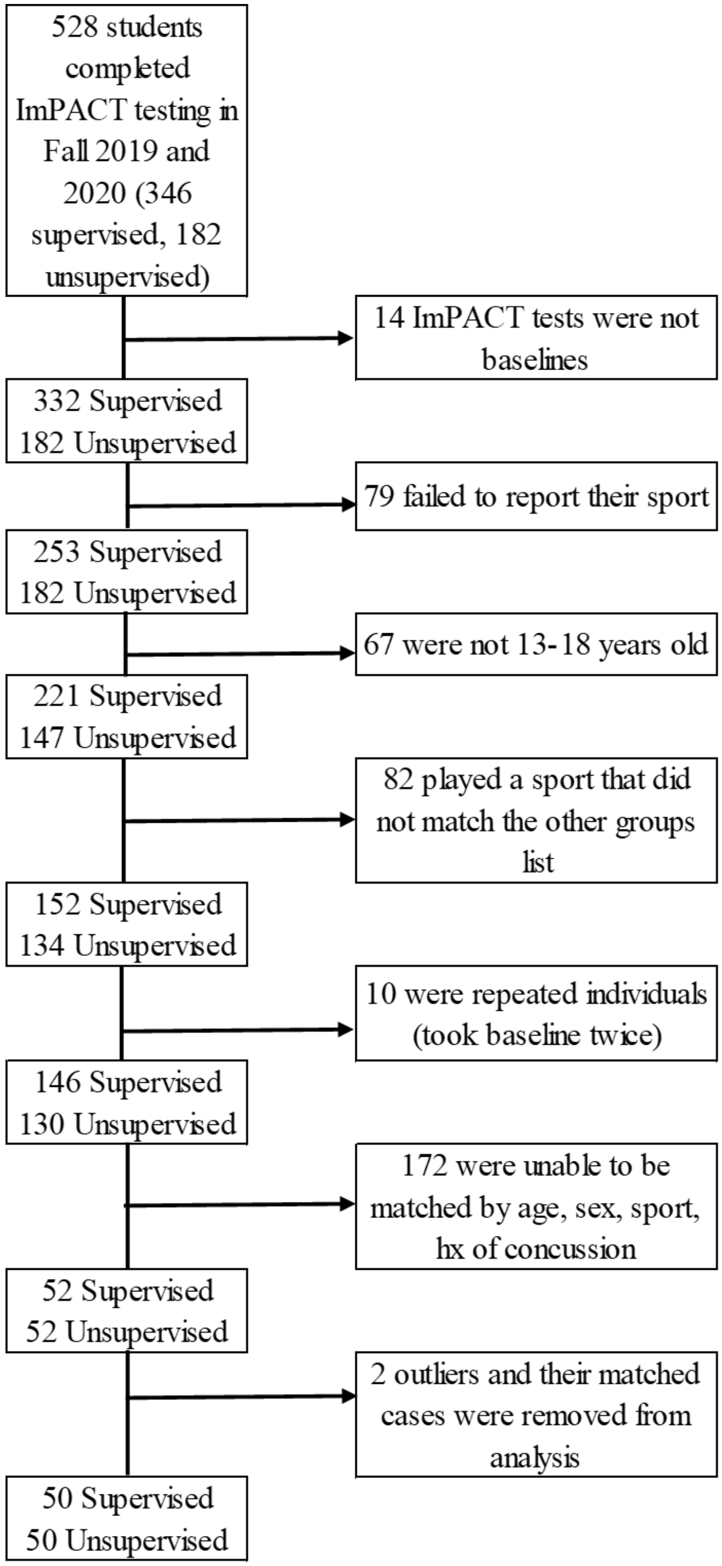


Figure 1. Flow chart illustrating the matching process of the at school (supervised) and at home (unsupervised) ImPACT baselines.

Demographics

Participants' ages ranged from 13 to 17 years old (14.5 ± 1.2 y/o), with 56% (n=56) of the participants being male. None of the participants reported a concussion history. There were a total of 12 sports, the most common being volleyball (n=12), tennis (n=8), and cross country (n=7) in each group. A full summary of the demographic's can be found in Table 1.

Table 1. Participants' Demographic Characteristics Total (n=100)

Characteristic	At School/Supervised	At Home/Unsupervised
Age, mean ± SD	14.5 ± 1.2	14.5 ± 1.2
Females, n (%)	22 (44)	22 (44)
Males, n (%)	28 (56)	28 (56)
Number of Concussions, mean ± SD	0 ± 0	0 ± 0
Sport, n (%)		
Volleyball	12 (24)	12 (24)
Tennis	8 (16)	8 (16)
Cross-Country	7 (14)	7 (14)
Soccer	6 (12)	6 (12)
Air Rifle	4 (8)	4 (8)
Baseball	3 (6)	3 (6)
Water Polo	3 (6)	3 (6)
Golf	2 (4)	2 (4)
Paddeling	2 (4)	2 (4)
Basketball	1 (2)	1 (2)
Softball	1 (2)	1 (2)
Wrestling	1 (2)	1 (2)

Composite Scores

In all the data analyzed, none of the baselines were flagged as invalid. Kolmogorov-Smirnov and Shapiro-Wilk tests were deemed significant for 3 out of 5 dependent variables (verbal memory, reaction time, impulse control), which indicates potential violation of normality assumption. However, since the total sample size of 100, in our study, was larger than the suggested G-power analysis[15] total sample size of 44, a parametric MANOVA was run for this data. The at home group had higher mean percentages in verbal and visual memory and better impulse control when compared to the supervised group. Verbal memory being 86.72 ± 9.24 for the self-administered and 84.28 ± 10.31 for the supervised group, visual memory 77.1 ± 12.85 and 76.28 ± 12.67, and impulse control 6.08 ± 4.88 and 6.48 ± 4.46, respectively. The unsupervised, self-administered group had worse visual motor speeds (36.1 ± 5.48, 37.89 ± 6.89) and slower reaction times (0.65 ± 0.1, 0.63 ± 0.08) when compared to the supervised groups. See Table 2 for the means and standard deviations.

Table 2. ImPACT Composite Scores by Group

	At School/Supervised	At Home/Unsupervised
Measure (unit)	mean ± SD	mean ± SD
Verbal memory (%)	84.28 ± 10.32	86.72 ± 9.24
Visual memory (%)	76.28 ± 12.67	77.10 ± 12.85
Visual motor speed (no.)	37.89 ± 6.89	36.10 ± 5.48
Reaction time (sec.)	0.63 ± 0.08	0.65 ± 0.10
Impulse control (no.)	6.48 ± 4.46	6.08 ± 4.88

Abbreviation: ImPACT, Immediate Post-Concussion Assessmentand Cognitive Testing.

Box Test, testing the assumption of homogeneity of variances, also deemed no significance between the two (*p*=0.18). Using Pillai's trace there was no significant difference in composite scores between groups [V = 0.45, F(5,94) = 0.89, *p* = .49]. See Table 3 for MANOVA results.

Table 3. Composite Score MANOVA Results

Effect	Test Statistic	Value	F	Error df	Sig. (p)	Partial η2	Noncentrality Parameter
Intercept	Pillai's Trace	0.996	5243.63	94.0	0	0.996	26218.15
	Wilk's Lambda	0.004	5243.63	94.0	0	0.996	26218.15
	Hotelling's Trace	278.920	5243.63	94.0	0	0.996	26218.15
	Roy's Largest Root	278.920	5243.63	94.0	0	0.996	26218.15
At School / At Home	Pillai's Trace	0.045	0.885	94.0	0.494	0.045	4.427
	Wilk's Lambda	0.955	0.885	94.0	0.494	0.045	4.427
	Hotelling's Trace	0.047	0.885	94.0	0.494	0.045	4.427
	Roy's Largest Root	0.047	0.885	94.0	0.494	0.045	4.427

Symptom Checklist

Independent T-test was conducted to compare scores for each symptom and total score. Levene's Test, assessment of the equality of variances, deemed 7 of the 22 symptoms as significant (fatigue, light sensitivity, numbness and tingling, feeling mentally foggy, difficulty concentrating, difficulty remembering, and visual problems). For these 7 symptoms, Independent T-test results looked at equal variances not assumed. Though not statistically significant, the supervised group had a higher means in the total symptom score when compared to the unsupervised, self-administered group (6.58±8.19 and 6.22±7.52 respectively, p=0.82). The supervised group also reported slightly higher symptom scores when compared to the unsupervised, self-administered group for headache (.36±.66, .34±.96, p=0.90), vomiting (.10±.36, .06±.31, p=0.56), sleeping less than usual (.80±1.13, .64±1.05, p=0.46), drowsiness (.18±.56, .10±.30, p=0.38), light sensitivity (.38±.83, .20±.50, p=0.19), noise sensitivity (.10±.36, .08±.34, p=0.78), irritability (.44±1.09, .32±.74, p=0.52), sadness (.68±1.0, .62±.92, p=0.76), numbness and tingling (.10±.36, .04±.20, p=0.31), feeling mentally foggy (.30±.65, .10±.51, p=0.09), and difficulty remembering (.46±.86, .20±.67, p=0.10). See Table 4 for the complete list of symptom scores means and standard deviations.

Table 4. ImPACT Symptom Scores

Symptom	Group: Supervised mean ± SD	Group: Unsupervised mean ± SD	95% Confidence Interval upper (lower)	Levene's Test *p* value
Headache	0.36 ± 0.66	0.34 ± 0.96	-0.31 (0.35)	0.93
Nausea	0.02 ± 0.14	0.02 ± 0.14	-0.06 (0.06)	1.00
Vomiting	0.10 ± 0.36	0.06 ± 0.31	-0.10 (0.18)	0.26
Blancing Problems	0.10 ± 0.30	0.14 ± 0.61	-0.23 (0.15)	0.33
Dizziness	0.14 ± 0.41	0.20 ± 0.57	-0.26 (0.14)	0.22
Fatigue	0.56 ± 0.93	0.86 ± 1.21	-0.73 (0.13)	0.04*
Trouble falling asleep	0.24 ± 0.56	0.26 ± 0.63	-0.26 (0.22)	0.63
Sleeping more than usual	0.16 ± 0.42	0.20 ± 0.42	-0.21 (0.13)	0.40
Sleeping less than usual	0.80 ± 1.13	0.64 ± 1.05	-0.27 (0.59)	0.39
Drowsiness	0.18 ± 0.56	0.10 ± 0.30	-0.10 (0.26)	0.07
Light Sensitivity	0.38 ± 0.83	0.20 ± 0.50	-0.09 (0.45)	0.01*
Noise Sensitivity	0.10 ± 0.36	0.08 ± 0.34	-0.12 (0.16)	0.59
Irritability	0.44 ± 1.09	0.32 ± 0.74	-0.25 (0.49)	0.17
Sadness	0.68 ± 1.00	0.62 ± 0.92	-0.32 (0.44)	0.26
Nervousness	0.24 ± 0.63	0.40 ± 0.86	-0.46 (0.14)	0.06
Feeling more emotional	0.34 ± 0.69	0.52 ± 1.04	-0.53 (0.17)	0.05
Numbness or tingling	0.10 ± 0.36	0.04 ± 0.20	-0.06 (0.18)	0.04*
Feeling slowed down	0.22 ± 0.58	0.22 ± 0.62	-0.24 (0.24)	0.93
Feeling mentally foggy	0.30 ± 0.65	0.10 ± 0.51	-0.03 (0.43)	0.003*
Difficulty concentrating	0.38 ± 0.67	0.64 ± 1.14	-0.63 (0.11)	0.01*
Diffiuculty remembering	0.46 ± 0.86	0.20 ± 0.67	-0.05 (0.57)	0.01*
Visual Problems	0.28 ± 0.70	0.06 ± 0.31	0.00 (0.44)	<0.01*
Total symptom score	6.58 ± 8.19	6.22 ± 7.52	-2.76 (3.48)	0.47

Abbreviation: ImPACT, Immediate Post-Concussion Assessmentand Cognitive Testing.
$*p < .05$

Of the 22 symptoms, Independent T-test showed that only visual problems had a significant difference in the supervised group when compared to the unsupervised individual group (.28±.79 supervised and .06±.31 unsupervised; *p*=.047). Table 5 has the complete list of T-test results.

Table 5. ImPACT Symptom T-test Results

Symptom	df	Mean Difference	Standard Error Difference	significance (2-tailed)
Headache	98.0	0.02	0.165	0.904
Nausea	98.0	0.00	0.028	1.00
Vomiting	98.0	0.04	0.068	0.558
Blancing Problems	98.0	-0.04	0.096	0.677
Dizziness	98.0	-0.06	0.099	0.546
Fatigue*	91.81	-0.3	0.216	0.168
Trouble falling asleep	98.0	-0.02	0.119	0.867
Sleeping more than usual	98.0	-0.04	0.087	0.648
Sleeping less than usual	98.0	0.16	0.217	0.463
Drowsiness	98.0	0.08	0.09	0.377
Light Sensitivity*	79.91	0.18	0.137	0.192
Noise Sensitivity	98.0	0.02	0.071	0.777
Irritability	98.0	0.12	0.186	0.521
Sadness	98.0	0.06	0.192	0.756
Nervousness	98.0	-0.16	0.15	0.289
Feeling more emotional	98.0	-0.18	0.176	0.308
Numbness or tingling*	75.63	0.06	0.059	0.309
Feeling slowed down	98.0	0.00	0.12	1.00
Feeling mentally foggy*	92.56	0.20	0.116	0.088
Difficulty concentrating*	79.07	-0.26	0.187	0.167
Diffuculty remembering*	92.37	0.26	0.154	0.096
Visual Problems*	67.86	0.22	0.109	0.047
Total symptom score	98.0	0.36	1.57	0.819

Abbreviation: ImPACT, Immediate Post-Concussion Assessmentand Cognitive Testing.
*Equal variances not assumed (Welch Method)

Discussion

There are few studies that compare different ways to administer ImPACT baseline tests. Of those studies, they compared either supervised and unsupervised[3] or group and individual[2 8 9] test taking. This is the first study to combine these two conditions and report the comparison of ImPACT baseline tests of supervised groups to those of unsupervised individually. Our study concluded that there were no significant differences in composite scores between the two groups; however, from a health care professional standpoint it would be best to keep test administration supervised. Supervision ensures that the test taker did in fact take the test for themselves and there are no outside factors contributing to potential distraction.

One observation to be made is Moser et al.[2] found that baselines taken in an individual setting had better visual motor speed and reaction time scores. That study concluded that group settings are subjected to higher rates of distraction and test interruptions than individual testing. The same two composite scores that Moser et al. deemed significant also had a significant difference in Kuhn et al.[3] study, but in favor of the supervised baselines. Kuhn et al. concluded that those who took the test unsupervised only thought that accuracy mattered and may not have considered that speed and reaction as factors that contribute to the final composite scores.

An observation noted in our study is that the groups had combined factors of supervision and group vs. individual setting and previous literature concluded that those two factors canceled each other out. Being those who took the test in an individual setting had better visual motor speed and RT and those who were not supervised had worse visual motor speed and RT, and vice versa. This can explain why there was no statistically significant differences between our studies groups in visual motor speed and reaction time. Though not statistically significant, our results showed similar trend of the unsupervised, self-administered individuals having slightly worse visual motor speed and slower reaction time means when compared to the supervised group.

Since this is the first study that combines two factors in each group, it is unsure which factor could have a bigger influence on the individual taking the test (group, individual, supervised, or unsupervised). It can be speculated that supervision has the greater influence on testing scores because our study showed that the unsupervised individual test takers had worse visual motor speeds and RTs when Moser et al.[2] found supervised individual testing had greater composite scores of the same two. A future study design that separately tests these four factors, using the same individuals with scripted instructions would be beneficial in testing which factor contributes more to test scores. This also leads into the importance of having a baseline that is tailored to that individual to ensure the most appropriate route for treating a possible head injury.[16 17]

Looking at the studies that compared group testing to individual testing, they all found and concluded different findings. Moser et al.[2], Vaughn et al.[9], and French et al.[8] were each cohort studies looking at the similar age ranges to the ones recorded in this study. Vaughn et al. found that those taken in groups had a higher error rate in impulse control and invalid baselines, Moser et al. found that baselines taken in an individual setting had better visual motor speed and reaction time scores, and French et al. found no differences. Both Vaughn et al. and French et al. concluded that regulated baseline administration and thorough instruction are key when distributing the ImPACT baseline and post-concussion tests. French et al. used specific instructions pre the ImPACT manual and encouraged test takers to read the instructions and do their best. Using an instructional script when administering ImPACT tests is a good clinical consideration to implement so there is no confusion for the test taker.

French et al.[8] found that baselines taken in a group setting reported lower symptom severity scores than those who took the baseline individually, while Moser et al.[2] found no differences between the two groups. Although not significant, the total symptom score means for unsupervised, at home baseline (6.22 ± 7.52) were less than those who took the baseline in a supervised group (6.58 ± 8.19). In our study, the symptom of visual problems was significantly higher in the supervised groups than the unsupervised at home. Since our supervised group was tested after school in the affiliated schools library, the prolonged exposure to fluorescent lighting could have been the influencing factor.[18]

Limitations

Due to this study being retrospective in nature, there were several uncontrollable factors to consider. Although the groups were matched, comparing NP tests of two different individuals is a limitation due to each person having their own unique neurocognitive ability. Another consideration is that though the 2019-2020 (supervised group) baselines were consistent in administration, the 2020-2021 (unsupervised, self-administered) can not report the same. The ImPACT manual recommends the use of an external mouse as well as a quiet environment with no distractions (ex. television, cellular device) and these were unknown in the self-administered group.[1] Another limitation is that most of the ‘at home’ participants “skipped” their demographic inputs. This means that the history of concussions and those with possible ADHD went unanswered. An ImPACT database search allowed to see if those participants who skipped had a previous concussion in the past; however, the ImPACT database only accounts for however long the participants played school affiliated sports. So, this does not account for any concussions sustained before enrolling in school affiliated sports. Though, it is shown in the literature that having a history of one or more concussions does not affect future ImPACT baselines.[19-21] Additionally, all participants are from a specific geographic location in the United States, so results may not be applicable to athletes in other parts of the country and in other countries. Future studies testing the same individual in each group with consistent instruction are warranted to better address supervised group scores to unsupervised individual scores.

Conclusion

Baseline testing in neuropsychological tests is a common and beneficial practice for concussion management; however, the best administrational practice for optimal effort scores remains unclear. This study was able to look at a group of athletes subjected to conditions put in place due to the COVID-19 pandemic restrictions. Although there were no statistical differences found in the composite scores and our directional hypothesis was rejected, other studies have concluded differences in group, individual, supervised, and un-supervised ImPACT baseline testing.[2 3 9] Our study can expose further questions on baseline administration in the future if past COVID-19 restrictions return or other circumstances arise. A study with consistent control and standardization of group and individual testing environments with scripted instructions to both supervised and unsupervised groups could help fill those gaps and present beneficial answers.

Review of Literature

According to the Centers for Disease Control and Prevention (CDC), concussion baseline testing is a good tool for trained health care professionals to use to compare with post-injury test scores and to aid in assistance when making diagnosis and treatment decisions.[22] ImPACT has also been found to be one of the most used concussion assessments.[23] Considering the pandemic of COVID-19 that is currently affecting sports across the world, the administration of ImPACT testing has changed and it is important to know if these changes could potentially influence a user's baseline scores. The purpose of this literature review was to cover all aspects of the ImPACT test and previous studies covering group and individual test-taking.

ImPACT

ImPACT is a computer-based NP test battery that uses objective measures of neurocognitive functioning to assist healthcare professionals in the assessment and management of concussions.[1] NP assessment is used to check the users performance in the core cognitive domains (the composite scores). The five cognitive domains include verbal memory, visual memory, visual motor speed, reaction time, and impulse control. Each domain is given a numerical score based on the module scores (further details of testing modules and scoring in Methods) in order to differentiate between normal cognitive function and abnormal cognitive function (mTBI, mental illness, or disease).[1 24] Normal cognitive function in this context refers to ImPACT's normative data.

Invalidity of the ImPACT Test

A survey of 399 ATs across the US revealed that 94% administer concussion baseline tests to their athletes.[25] Baselines are used to help compare post-concussion test scores to the athletes personalized pre-concussion normative. Of the 94% of ATs who administered baselines, only 54% of them checked for invalid tests. Checking for invalid tests means the AT is checking the composite scores to see if the users scores were below the recommended ImPACT scores. An invalid test could suggest a possible mTBI but it could indicate possible sandbagging or pre-existing factors.[1 4-6] Sandbagging is when the test taker purposely tries to score low on their baseline so they could return back to sport early, post-mTBI. Athletes who attempt to sandbag can influence the rate of validity.

Due to the threat of sandbagging, ImPACT has established five “red flags” to catch those purposely trying to score low. These five “red-flags” are based off of the composite scores and flagged when the score is above or below ImPACTs recommendations.[7 26] Along with ImPACTs “red flags,” ImPACT also has Invalidity Indicators. An invalid ImPACT score signifies that the test taker scored below the ImPACT recommended invalidity indicators for one of the scores pulled from select modules. An example would be X’s and O’s Total Incorrect > 30 or Impulse Control Composite > 30, etc. An invalid score can still occur for individuals with no ongoing mTBI or other neurocognitive disability. This can happen through pre-existing factors (sex, treatment of headache/migraine, mental illness, and ADHD diagnosis)[4-6], sandbagging, and lack of motivation to do well.[1]

A systematic review[12] found that in three of the 12 studies reviewed, ImPACT failed to detect 20% of coached sandbagging. This means that the test takers were given instructions on how to score low without having an invalid test score and 20% were able to do so. In the same systematic review, the 12 articles assessed had included prevalence rates in their studies and found ImPACT to have a prevalence rate of invalid performance on baseline testing to range between 2.7-27.9%.[1-3 5 27 28] After removing an outlier, the weighted prevalence rate came to be 6.1% of invalid tests. When looking at the studies that tested youth and high school students (10-18 y/o), the weighted prevalence rate came to be 6.0%. This means that 6% of individuals who are 10-18 years old who are administered ImPACT will receive an invalid score.[12] A retrospective study[26] of 6,346 high school athletes in Hawaiʻi found 51.99% had valid baseline profiles, 4.24% had invalid scores, and 47.42% had red flag (possible sandbag) scores and 3.64% had both invalid and red flag. Of the 47.42% who obtained sandbagging scores: 52.94% had one sandbagging score, 26.65% had two sandbagging scores, and 9.7% had three sandbagging scores. Most articles pertaining to sandbagging had instructed their participants to purposely score low for the sake of figuring out validity.

Though Tsushima et al[26] found high rates of invalid scores and sandbagging, Schatz et al[29] found differing results. Schatz et al compared three groups of college aged students. The three groups were separated by either being coached on how to successfully sandbag, told to fake a concussion without being coached, or to do the best they could. It found that ImPACT was able to identify 95% of the naïve (non-coached) and 100% of the coached test-takers. The ImPACT manual[1] talked about a study that looked at 75 undergraduate collegiate athletes who had already been administered their baselines and were re-administered a second time and instructed to "fake bad" without reaching the validity indicators. In this study only eight of the 75 were able to fake the test successfully, meaning that ImPACT had an 89% successful sandbagging detection rate. Conversely, several authors noted that the occurrence of sandbagging is rare.[5 28]

ImPACT Reliability

A systematic review[30] of test-retest reliability for the ImPACT test reviewed nine studies that consisted of mostly high school and college age participants, one study included a professional hockey team. The test-retest time frame spanned from 24 hours to two years. Baumgartner et al[31] Intraclass Correlation Coefficient (ICC) guidelines were used in this systematic review. The ICC guidelines are used to rate the reliability and they suggest when a coefficient exceeds 0.8 it has good reliability, moderate reliability occurs when a coefficient is between 0.6 and 0.79, and poor reliability occurs when the coefficient falls below 0.6. The ICC for each composite score results in Alsalaheen et al[30] were the following: verbal memory (0.23-0.79), visual memory (0.26-0.85), processing speed (0.38-0.91), and reaction time (0.39-0.88). The review also sums up the percentage of participants who had a significant change in scores from their retest. These percentages were: verbal memory (5-26.8%), visual memory (2.2-19.6%), visual motor processing speed (4-24%), and reaction time (4-23.2%). Across all the reviewed studies, ImPACT reliability for all composite scores fell between low and moderate. A 2017 meta-analysis[13] concluded with the same results. This means ImPACT does not achieve good reliability. It is an ongoing debate of ImPACT's reliability and validity.[32]

Sensitivity & Specificity for ImPACT

Sensitivity, for ImPACT, measures the amount of 'positives' that are identified correctly. Meaning, the ability of ImPACT to correctly identify those with mTBI. While Specificity is the opposite, the ability to correctly identify those who are healthy (no mTBI). The manual reviews four studies from 2006 to 2015 who covers this topic.[1] Across the four studies, 381 subjects were looked at and the authors of the manual concluded that ImPACT has the ability to yield 91.4% sensitivity and 69.1% specificity for those who were symptomatic (reporting symptoms). For those who were asymptomatic (reporting no symptoms), and suspected of hiding a possible concussion, the authors concluded that ImPACT yields 94.6% sensitivity and 97.3% specificity.[1]

Pre-Existing Factors

Baseline scores can be determined by several pre-existing factors, such as age, sex, previous concussions, and the presence of ADHD. A cross-section study[5] of 504 athletes (10-18 y/o) found that younger (10-12 y/o) test takers tend to have a higher rate of invalid baselines (7%) compared to older individuals (13-18 y/o) with a rate of 2.7%. A retrospective study[32] of 5741 adolescent athletes in Hawai‘i found that ages 13 to 15 had lower composite scores than 16 to 18 year-olds. Though there were consistent findings for age playing a factor in baseline scores, some studies had conflicting findings when comparing males to females. For example, Tsushima et al[32] reported females to score higher in visual motor speed, reaction time, impulse control, and higher total symptom scores than males. Males had high scores in visual memory. A cross-sectional study[6] of 486 division I college athletes found that sex had no difference in scores but a systematic review[12] found males to have higher rates of invalid performances when compared to females.[2 5 33]

A systematic review and meta-analysis[19] found that adolescents with a history of one previous concussion have lower baseline scores in visual memory than those with no previous concussion history. There were no differences found in those with more than one previous concussion. Tsushima et al[34] compared 41 high school athletes in Hawai‘i, who had not received a sport related concussion (SRC) between 2012-2013 school year, and 39 athletes who had received a SRC. Among those who had sustained one SRC, five reported a history of ADHD, one reported a history of special education, and four reported a history of prior SRCs. The group with no mTBI only had one person with ADHD and none reported prior SRCs. He found in the group who sustained a SRC no differences in ImPACT scores pre and post-concussion. A retrospective study of 483 high school athletes was put into three groups: no previous concussion (409), one previous concussion (58), and two previous concussions (16). There was no significant difference found between the three groups ImPACT composite scores.[21]

The normative data sample ImPACT conducted and currently uses (more details below) found that those who reported having a LD and ADHD had a statistically lower score in their visual motor and reaction time composite scores compared to those who reported having no LD or ADHD. Cottle et al[6] cross-sectional study of 486 division I college athletes found those with ADHD tend to have lower composite scores and a higher rate of invalid baselines than those without. Another cross-sectional study[5] of 502 adolescent athletes (10-18 y/o) had the same conclusion. Cottle et al[6] also found that those with chronic headache/migraine and those with a mental illness (depression, bipolar, etc.) will have higher symptom scores.

ImPACT's Normative Data

If a user does not have a baseline and only takes a post-concussion ImPACT test, the users scores will be compared to ImPACTs normative data that was collected and averaged out. From 2006-2007, ImPACT collected 16,566 test scores from individuals aged 12 to 59 years old; the averages taken from this sample is the set standardization ImPACT uses as their normative data. The 16,566 tests were taken from high schools affiliated with Western Pennsylvania Interscholastic Athletic League (WPICAL), colleges with previous experience with ImPACT across America, Professional sports teams across American, and coaches, teachers, and school administrators from the affiliated schools. All the tests were administered, supervised, and collected by neuropsychologists, psychologists, graduate students majoring in athletic training, neuropsychology, or psychology, ATs, and nurses who were properly trained in the administration and use of ImPACT. Prior to subject inclusion, all participants had to complete a physical examination, they were all also symptom free and reported no occurring medical or psychological conditions. Subjects also reported no previous history of meningitis, epilepsy, or any other neurological diseases.[1]

Of the 16,566 participants, 72% were male and 28% female. Though the percentage of females were greatly lower than males, the sample size was sufficient for establishing normative data.[1] The normative sample consisted of athletes who participated in the following sports: male and female soccer, lacrosse, swimming/diving, rowing, volleyball, track and field, and cross country. Sports that only males participated in included football wrestling, and baseball; only female sports included softball, cheerleading, and field hockey.

Most ages were grouped together in an age range of 12, 13-15, 16-18, 19-29, 30-39, 40-49, and 50-59 years old. The 13-15 and 16-18 ages made up most of the 16,566 participants; both groups containing 38% of the sample size (76% of the total). Less than one percent of the sample size reported any LD and 4% reported a history of ADHD. If there is no baseline for a user and only post-concussion results, the manual states that the normative values will show in the clinical report for clinicians to use to compare.[1]

The United States Census Bureau shows that over 75% of the population is white.[35] So, it can be suggested that the normative scores for ImPACT can overlook the norms of minorities. In a retrospective study[36] of ImPACT baselines of 405 professional baseball athletes (304 players were native-English speakers and 101were native Spanish speakers) it showed that native English speakers score higher verbal memory, visual memory, visual motor speed, reaction time, and total symptom scores than native Spanish speakers. Blake et al[37] had 58 bilingual English–

Spanish-speaking undergraduate students take an ImPACT baseline in both English and Spanish. The results showed that Spanish-English speakers scored better on ImPACT when it was in English compared to Spanish. A prospective case–control[38] design matched two cohorts (age, sex, and concussion history) of 48 white and 48 black high school and college aged athletes and found no difference in scores between the two.

Tsushima et al[39] looked at 247 private high school student athletes in Hawai‘i and found that the male athletes scored similar composite scores to ImPACTs suggested normative. However, the retrospective study[32] of 5741 high school athletes (mentioned above) from 36 schools in Hawai‘i found that athletes of asian backgrounds scored higher in all 6 modules than those of hawaiian/pacific islander descent. Athletes of asian descent also scored higher in visual motor speed, reaction time and impulse control than those of a mixed ethnical background. Those of a mixed ethnical background had better scores all around (except impulse control) than those of hawaiian/pacific islander backgrounds. In the same study, Tsushima et al found that high school athletes who participated in soccer, basketball, and volleyball had better baseline scores than those who participated in football.[32]

Testing Environment

ImPACT baseline is usually administered in a group setting for those playing on a school affiliated sports team due to convenience and time management. Post-concussion testing is usually administered individually. Moser et al[2] performed a cohort study comparing the test results of high school athletes who took the ImPACT baseline either in a group or individually. Of the two groups, the ‘group setting’ consisted of 164 participants and the ‘individual setting’ had 167 participants. Moser et al found more invalid baselines occurred to those taken in a supervised group setting than to those taken in a supervised individual setting. Also, tests taken in a group had worse motor processing speed and reaction time than the individual test takers. However, the proctors did not have a set script for delivering thorough instructions. Another cohort study[9] comparing 313 individuals who were tested individually to 626 individuals tested in a group setting (ages 5-18 years) found that 11.2% of baselines taken in a supervised, instructed group were invalid compared to 1.7% invalid baselines in the individually supervised, instructed group. Findings also showed those taken in a group had a higher error rate on the impulse control score.

French et al[8] argued that it did not matter if the baseline was taken in a group or individual setting, the only thing that mattered was if the test takers had clear and specific instructions. French et al was a cohort study looking at 500 athletes testing in a group environment and 500 athletes testing in an individual environment (all aged 10-18 years). The results found no difference between group and individual baselines because both were administered scripted instructions before proceeding with the test.[8] Kuhn et al[3] is the only study that looked at 1070 supervised vs 1070 self-administered, non-supervised matched (number of prior concussions, sex, and grade) baselines of high school athletes. The results found that supervised baselines had higher visual motor speeds and faster reaction times. However, it does not mention if either group was given instructions prior to administering the baseline tests.

Appendices

Appendix A: ImPACT PCSS

ImPACT
APPLICATIONS, INC.
ConcussionManagement.com

Post-Concussion Symptom Scale

Patient's Name: ____________________

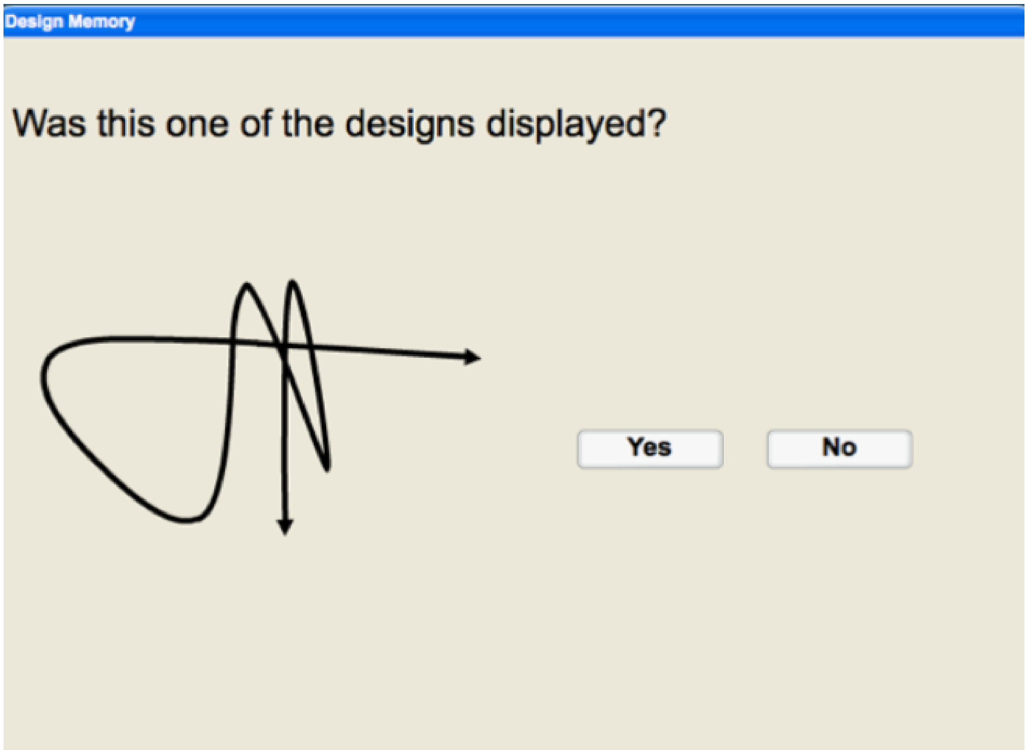
Date of Birth: ____________

Please use the following scale to rate each symptom:

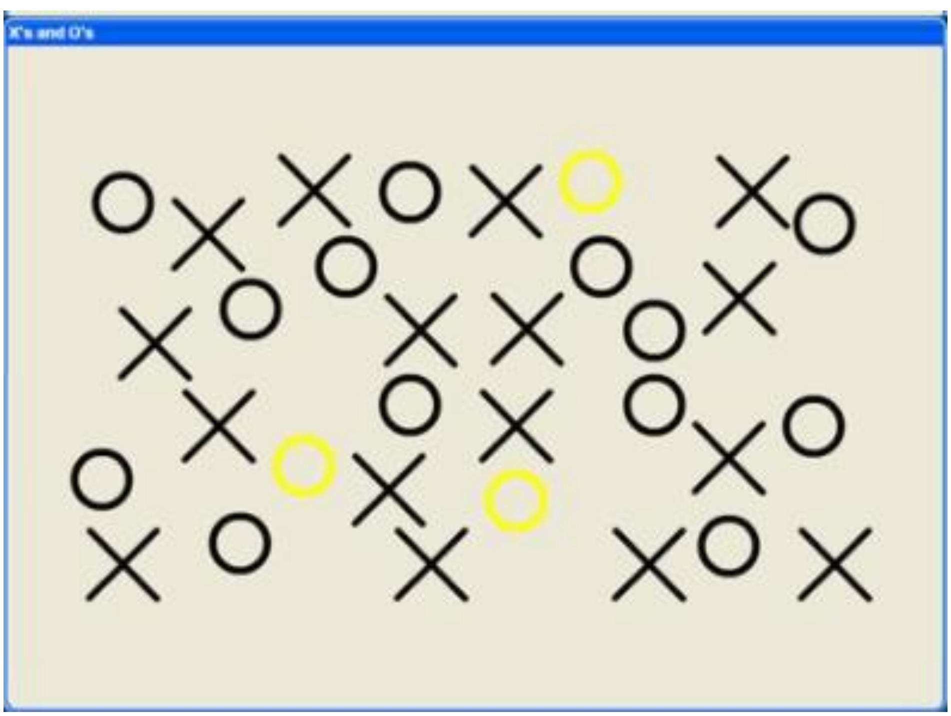
None		Mild		Moderate		Severe
0	1	2	3	4	5	6

SYMPTOMS	SEVERITY RATING						
	Date:	Date:	Date:	Date:	Date:	Date:	Date:
Headache							
Nausea							
Vomiting							
Balance Problems							
Dizziness (spinning or movement sensation)							
Lightheadedness							
Fatigue							
Trouble falling asleep							
Sleeping more than usual							
Sleeping less than usual							
Drowsiness							
Sensitivity to light							
Sensitivity to noise							
Irritability							
Sadness							
Nervous/ Anxious							
Feeling more emotional							
Numbness or tingling							
Feeling slowed down							
Feeling like "in a fog"							
Difficulty concentrating							
Difficulty remembering							
Visual problems							
Other							
Total							

Appendix B: Design Memory Module



Appendix C: X's and O's Module



Appendix D: Visual of Module Scoring Components

DESIGN MEMORY

Hits (immediate)					
Correct distractors (immed.)					
Learning percent correct					
Hits (delay)					
Correct distractors (delay)					
Delayed memory % correct					
Total percent correct					

X's AND O's

Total correct (memory)					
Total correct (interference)					
Average correct RT (interference)					
Total incorrect (interference)					
Average incorrect RT (interference)					

SYMBOL MATCH

Total correct (visible)					
Average correct RT (visible)					
Total correct (hidden)					
Average correct RT (hidden)					

COLOR MATCH

Total correct					
Average correct RT					
Total commissions					
Average commissions RT					

THREE LETTERS

Total sequence correct					
Total letters correct					
% of total letters correct					
Average time to first click					
Average counted					
Average counted correctly					

Appendix E: Manual Example of Visual Memory Composite Scoring

Visual Memory Composite

Evaluates visual attention and scanning, learning, and memory

This score in its current form is comprised of the average of:

- Design Memory (module 2) Total Percent Correct
- X's and O's (module 3) (Total Correct Memory)/12*100

Design Memory		
Hits (Immediate)	12	10
Correct distractors (immed.)	10	5
Learning percent correct	92%	63%
Hits (delay)	12	4
Correct distractors (delay)	9	7
Delayed memory pct. correct	88%	46%
Total percent correct	90%	54.5%

X's and O's		
Total correct (memory)	8	5
Total correct (interference)	109	67
Avg. correct RT (interference)	0.55	1.17
Total correct (interference)	3	3
Avg. incorrect RT (interfer.)	0.43	1.04

EXAMPLE	
Design Memory Total Percent Correct =	90 %
X's and O's (Total Correct Memory)/12= 8/12*100	66.7%
Total Divided by 2 = 156.7/2 =	78

Visual Memory Composite Score= 78

References

1. Lovell MR. *ImPACT Test: Administration and interpretation manual*. Pittsburgh, PA: ImPACT Applications, Inc., 2016.
2. Moser RS, Schatz P, Neidzwski K, Ott SD. Group versus individual administration affects baseline neurocognitive test performance. Am J Sports Med 2011;**39**(11):2325-30 doi: 10.1177/0363546511417114[published Online First: Epub Date]|.
3. Kuhn AW, Solomon GS. Supervision and computerized neurocognitive baseline test performance in high school athletes: an initial investigation. J Athl Train 2014;**49**(6):800-5 doi: 10.4085/1062-6050-49.3.66[published Online First: Epub Date]|.
4. Nelson LD, LaRoche AA, Pfaller AY, et al. Prospective, Head-to-Head Study of Three Computerized Neurocognitive Assessment Tools (CNTs): Reliability and Validity for the Assessment of Sport-Related Concussion. J Int Neuropsychol Soc 2016;**22**(1):24-37 doi: 10.1017/S1355617715001101[published Online First: Epub Date]|.
5. Lichtenstein JD, Moser RS, Schatz P. Age and test setting affect the prevalence of invalid baseline scores on neurocognitive tests. Am J Sports Med 2014;**42**(2):479-84 doi: 10.1177/0363546513509225[published Online First: Epub Date]|.
6. Cottle JE, Hall EE, Patel K, Barnes KP, Ketcham CJ. Concussion Baseline Testing: Preexisting Factors, Symptoms, and Neurocognitive Performance. J Athl Train 2017;**52**(2):77-81 doi: 10.4085/1062-6050-51.12.21[published Online First: Epub Date]|.
7. Lovell MR. Clinical Interpretation Manual, 2007.
8. French J, Huber P, McShane J, Holland CL, Elbin RJ, Kontos AP. Influence of Test Environment, Age, Sex, and Sport on Baseline Computerized Neurocognitive Test Performance. Am J Sports Med 2019;**47**(13):3263-69 doi: 10.1177/0363546519875137[published Online First: Epub Date]|.
9. Vaughan CG, Gerst EH, Sady MD, Newman JB, Gioia GA. The Relation Between Testing Environment and Baseline Performance in Child and Adolescent Concussion Assessment. Am J Sports Med 2014;**42**(7):1716-23 doi: 10.1177/0363546514531732[published Online First: Epub Date]|.
10. DOE H. Return to Learn. Secondary Return to Learn December 2, 2020 2020. http://www.hawaiipublicschools.org/ConnectWithUs/MediaRoom/PressReleases/Pages/school-year-2020-21.aspx.
11. Association HHSA. HHSAA Updates on COVID-19. Secondary HHSAA Updates on COVID-19 2020. https://www.sportshigh.com/resources/COVID-19.
12. Gaudet CE, Weyandt LL. Immediate Post-Concussion and Cognitive Testing (ImPACT): a systematic review of the prevalence and assessment of invalid performance. Clin Neuropsychol 2017;**31**(1):43-58 doi: 10.1080/13854046.2016.1220622[published Online First: Epub Date]|.
13. Farnsworth JL, 2nd, Dargo L, Ragan BG, Kang M. Reliability of Computerized Neurocognitive Tests for Concussion Assessment: A Meta-Analysis. J Athl Train 2017;**52**(9):826-33 doi: 10.4085/1062-6050-52.6.03[published Online First: Epub Date]|.
14. Alsalaheen B, Stockdale K, Pechumer D, Broglio SP. Validity of the Immediate Post Concussion Assessment and Cognitive Testing (ImPACT). Sports Med 2016;**46**(10):1487-501 doi: 10.1007/s40279-016-0532-y[published Online First: Epub Date]|.

15. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009;**41**(4):1149-60 doi: 10.3758/BRM.41.4.1149[published Online First: Epub Date]|.
16. Hinton-Bayre AD. Normative versus baseline paradigms for detecting neuropsychological impairment following sports-related concussion. Brain Impairment 2015;**16(2)**:80-89
17. Louey AG, Cromer JA, Schembri AJ, et al. Detecting cognitive impairment after concussion: sensitivity of change from baseline and normative data methods using the CogSport/Axon cognitive test battery. Arch Clin Neuropsychol 2014;**29**(5):432-41 doi: 10.1093/arclin/acu020[published Online First: Epub Date]|.
18. <Wilkins 1990- A tint to reduce eye-strain from fluorescent lighting.pdf>.
19. Alsalaheen B, Stockdale K, Pechumer D, Giessing A, He X, Broglio SP. Cumulative Effects of Concussion History on Baseline Computerized Neurocognitive Test Scores: Systematic Review and Meta-analysis. Sports Health 2017;**9**(4):324-32 doi: 10.1177/1941738117713974[published Online First: Epub Date]|.
20. Tsushima WT, Jordan A, Tsushima VG, Murata NM. Effects of a Single Concussion During the School Year on the Academic Performance and Neuropsychological Functioning of High School Athletes. Hawaii J Health Soc Welf 2020;**79**(7):212-16
21. Tsushima WT, Geling O, Arnold M, Oshiro R. Effects of Two Concussions on the Neuropsychological Functioning and Symptom Reporting of High School Athletes. Appl Neuropsychol Child 2016;**5**(1):9-13 doi: 10.1080/21622965.2014.902762[published Online First: Epub Date]|.
22. Prevention CfDCa. FAQs about Baseline Testing. Secondary FAQs about Baseline Testing 2015. https://www.cdc.gov/headsup/basics/baseline_testing.html#:~:text=Baseline%20tests%20are%20used%20to,presence%20of%20any%20concussion%20symptoms.
23. Dessy AM, Yuk FJ, Maniya AY, et al. Review of Assessment Scales for Diagnosing and Monitoring Sports-related Concussion. Cureus 2017;**9**(12):e1922 doi: 10.7759/cureus.1922[published Online First: Epub Date]|.
24. McCrory P, Meeuwisse WH, Aubry M, et al. Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. Br J Sports Med 2013;**47**(5):250-8 doi: 10.1136/bjsports-2013-092313[published Online First: Epub Date]|.
25. <Covassin 2009- Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) Practices of Sports Medicine Professionals.pdf>.
26. Tsushima WT, Yamamoto MH, Ahn HJ, Siu AM, Choi SY, Murata NM. Invalid Baseline Testing with ImPACT: Does Sandbagging Occur with High School Athletes? Appl Neuropsychol Child 2019:1-10 doi: 10.1080/21622965.2019.1642202[published Online First: Epub Date]|.
27. Brooks BL, McKay CD, Mrazik M, Barlow KM, Meeuwisse WH, Emery CA. Subjective, but not objective, lingering effects of multiple past concussions in adolescents. J Neurotrauma 2013;**30**(17):1469-75 doi: 10.1089/neu.2012.2720[published Online First: Epub Date]|.
28. Nelson LD, Pfaller AY, Rein LE, McCrea MA. Rates and Predictors of Invalid Baseline Test Performance in High School and Collegiate Athletes for 3 Computerized Neurocognitive

Tests: ANAM, Axon Sports, and ImPACT. Am J Sports Med 2015;**43**(8):2018-26 doi: 10.1177/0363546515587714[published Online First: Epub Date]|.
29. Schatz P, Sandel N. Sensitivity and specificity of the online version of ImPACT in high school and collegiate athletes. Am J Sports Med 2013;**41**(2):321-6 doi: 10.1177/0363546512466038[published Online First: Epub Date]|.
30. Alsalaheen B, Stockdale K, Pechumer D, Broglio SP. Measurement Error in the Immediate Postconcussion Assessment and Cognitive Testing (ImPACT): Systematic Review. J Head Trauma Rehabil 2016;**31**(4):242-51 doi: 10.1097/HTR.0000000000000175[published Online First: Epub Date]|.
31. Baumgartner T, Jackson A, Mahar M, Rowe M. *Measurement for Evaluation in Physical Education and Exercise Science*. Dubuque, IA: McGraw-Hill, 1999.
32. Tsushima WT, Tsushima VG, Murata NM. ImPACT Normative Data of Ethnically Diverse Adolescent Athletes. Clin J Sport Med 2020;**30**(1):52-59 doi: 10.1097/JSM.0000000000000567[published Online First: Epub Date]|.
33. Wojtowocz M, Iverson G, Resch J, et al. Factors associated with invalid scores on baseline neurocognitive testing in athletes [Poster Abstract]. Archives of Clinical Neuropsychology, 2015:505.
34. WT WT, Jordan A, Tsushima V, Murata N. Effects of a Single Concussion During the School Year on the Academic Performance and Neuropsychological Functioning of High School Athletes. Hawaii J Health Soc Welf 2020;**79**(7):212-16
35. United States Census Bureau Quick Facts. Secondary United States Census Bureau Quick Facts 2019. https://www.census.gov/quickfacts/fact/table/US/PST045219.
36. Jones N, Walter K, R RC, Wright D, Raasch W, Young C. Effect of education and language on baseline concussion screening tests in professional baseball players. Clin J Sport Med 2014;**24**(4):284-8
37. Lehman Blake M, Ott S, Villanyi E, Kazhuro K, Schatz P. Influence of Language of Administration on ImPACT Performance by Bilingual Spanish-English College Students. Arch Clin Neuropsychol 2015;**30**(4):302-9 doi: 10.1093/arclin/acv021[published Online First: Epub Date]|.
38. Kontos AP, Elbin RJ, 3rd, Covassin T, Larson E. Exploring differences in computerized neurocognitive concussion testing between African American and White athletes. Arch Clin Neuropsychol 2010;**25**(8):734-44 doi: 10.1093/arclin/acq068[published Online First: Epub Date]|.
39. Tsushima W, Siu A. Neuropsychological Test Performance of Hawai‘i High School Athletes: Updated Hawai‘i Immediate Post-Concussion Assessment and Cognitive Testing Data. Hawaii J Health Soc Welf 2014;**73**(7):208-11